

بررسی ۳۰ شاخص کیفی اعتبار ساختاری و محتوایی آزمون‌های نظری دندانپزشکی

مریم بهاروند*، محمد حسین زاده**، زهرا جابری انصاری***، الهام عباس زاده****، حامد مرتضوی*#

* دانشیار گروه بیماری‌های دهان، دانشکده دندانپزشکی، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران

** پزشک، دانشجوی دوره PhD دانشکده آموزش علوم پزشکی، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران

*** دانشیار گروه ترمیمی و زیبایی، دانشکده دندانپزشکی، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران

**** دستیار تخصصی گروه بیماری‌های دهان، دانشکده دندانپزشکی دانشگاه علوم پزشکی کرمان، ایران

تاریخ ارائه مقاله: ۹۲/۱۰/۳۰ - تاریخ پذیرش: ۹۳/۶/۱۲

Evaluation of 30 Structural Validity and Content Quality Indices of Theoretical Dental Exams

Maryam Baharvand*, Mohammad Hoseinzadeh**, Zahra Jaberiansari***, Elham Abbaszadeh****, Hamed Mortazavi*#

* DDS, MS, Associate Professor, Dept of Oral Medicine, School of Dentistry, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

** MD, PhD Student, School of Medical Education, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

*** DDS, MS, Associate Professor, Dept of Operative Dentistry, School of Dentistry, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

**** DDS, Postgraduate Student, Dept of Oral Medicine, School of Dentistry, Kerman University of Medical Sciences, Kerman, Iran.

Received: 20 January 2014 ; Accepted: 3 September 2014

Introduction: Exam is an inseparable part of education and the main tool to evaluate the results of educational process. If based on scientific rules, it can show the level of achievement to educational goals. This study was accomplished to evaluate the quality of theoretical exams in different departments of Dental School, Shahid Beheshti University of Medical Sciences

Materials & Methods: In this cross-sectional study, the quality of all written final exams of the first theoretical courses of 10 departments were evaluated in terms of 30 indices regarding structural validity, content relevance and content coverage, differential index, difficulty index, value of distracters, and taxonomy. Millman's check list, content analysis table and standard formula were used to collect the data. SPSS software version 15 was used to analyze the data.

Results: Out of 544 questions, 67.5% were multiple choice questions, 13% were descriptive and 19.5% were true-false queries. Content relevance and content coverage was considered in 85.4% and 84.5% of questions respectively. Standard structural rules were observed in 96.4% of questions and 72.2% of exams. Difficulty index, differential index, and value of distracters were estimated to be 67.8%, 22.3%, and 24.5% respectively. 84.6% of questions were of taxonomy I, and the remaining 15.4% were of taxonomy II.

Conclusion: Content relevance, content coverage, consideration of structural rules, and difficulty index were acceptable in theoretical written exams, whereas taxonomy, differential index and value of distracters needed to be reconsidered.

Key words: Exam, content relevance, content coverage, structural validity.

Corresponding Author: hamedmortazavi2013@gmail.com

J Mash Dent Sch 2015; 38(4): 291-302 .

چکیده

مقدمه: آزمون جزء جدایی ناپذیر آموزش و ابزار اصلی جهت سنجش نتایج فرآیند آموزش می باشد. اگر آزمون بر اساس اصول علمی انجام گیرد می تواند نشانگر میزان دستیابی به اهداف آموزشی باشد. این مطالعه باهدف ارزیابی کیفیت آزمون های نظری بخش های مختلف دانشکده دندانپزشکی دانشگاه شهید بهشتی انجام گرفت.

مواد و روش ها: در این پژوهش توصیفی مقطعی، کلیه سؤالات آزمون های کتبی دروس ۱ نظری ۱۰ بخش از لحاظ ۳۰ شاخص مربوط به اعتبار ساختاری، ارتباط و پوشش محتوایی، ضریب تمیز، ضریب دشواری، ارزش مورد نظر برای هر گزینه انحرافی و تاکسونومی مورد بررسی قرار گرفتند. ابزار مورد استفاده جهت گردآوری اطلاعات چک لیست Millman، فرمول های استاندارد و جدول آنالیز محتوا بوده است. داده ها با استفاده از نرم افزار آماری SPSS نسخه ۱۵ تجزیه و تحلیل شدند.

یافته ها: از ۵۴۴ سوال بررسی شده، ۶۷/۵ درصد چند گزینه ای، ۱۳ درصد تشریحی و ۱۹/۵ درصد صحیح/ غلط بوده است. ارتباط محتوایی آزمون های مورد بررسی ۸۵/۴ درصد و پوشش محتوایی ۸۴/۵ درصد بوده است. همچنین قواعد ساختاری سوال در ۹۶/۴ درصد موارد و قواعد نگارشی آزمون در ۷۲/۲ درصد موارد رعایت شده بود. ضریب دشواری سوالات ۶۷/۸ درصد، ضریب تمیز ۲۲/۳ درصد و ارزش مورد نظر برای هر گزینه انحرافی ۲۴/۵ درصد برآورد شد. ۸۴/۶ درصد سوالات دارای تاکسونومی ۱ و ۱۵/۴ درصد سوالات، دارای تاکسونومی ۲ بودند.

نتیجه گیری: آزمون های کتبی دروس نظری برگزار شده در دانشکده دندانپزشکی شهید بهشتی از لحاظ ارتباط و پوشش محتوایی، رعایت قواعد ساختاری و ضریب دشواری در سطح مناسب قرار دارند ولی از لحاظ تاکسونومی، ضریب تمیز و ارزش مورد نظر برای هر گزینه انحرافی نیاز به بازنگری دارند.

کلمات کلیدی: آزمون، ارتباط محتوایی، پوشش محتوایی، اعتبار ساختاری. مجله دانشکده دندانپزشکی مشهد / سال ۱۳۹۳ دوره ۳۸ / شماره ۴: ۳۰۲-۲۹۱.

مقدمه

جای گذاشته و خطر به هدر رفتن تلاش استادان و سیستم آموزشی را خواهد داشت. لذا بررسی کیفیت سوالات و استاندارد بودن آزمون ضروری است.^(۲) با وجود این سطح از اهمیت، مطالعات جامعی که سوالات آزمون های نظری را از نظر شاخص های مختلف کیفی بررسی کرده باشند، اندک می باشد. نتایج حاصل از تحقیقات محدود صورت گرفته در این حوزه نشان می دهد که در اغلب آزمون ها از سوالات مناسب کمتر استفاده شده است و استادان مهارت کافی در ساخت آزمون های درسی را نداشته و برای اندازه گیری پیشرفت تحصیلی دانشجویان اغلب مفاهیم ساده و سطحی درس را اندازه گیری می کنند و در اکثر مواقع سوالات مبهم بوده و بسیاری از مدرسان بدون این که آموزش کافی در سنجش و اندازه گیری و شیوه ساخت آزمون ها کسب کرده باشند تدریس می نمایند.^(۳-۵) از آنجایی که برای رفع این مشکل ابتدا باید اولویت ها، کاستی ها و نقاط ضعف موجود در هر

ارزیابی عبارت است از فرایند نظام مند گردآوری، تحلیل و تفسیر داده ها به منظور ارزش گذاری و قضاوت؛ به طوری که تعیین کند آیا هدف مورد نظر تحقق یافته یا در حال تحقق یافتن است و میزان تحقق یافتن آن به چه میزان است.^(۱) دراین میان آزمون به عنوان ابزار اصلی ارزیابی، بخش جدایی ناپذیر فرآیند آموزش و تدریس بوده و چنانچه بر اساس اصول و موازین علمی و به طور مستمر انجام گیرد، نشانگر میزان دستیابی استاد و دانشجو به اهداف از پیش تعیین شده و یادگیری دانشجویان و در نهایت میزان دستیابی به اهداف آموزشی می باشد. این در حالی است که اگر یک آزمون از طراحی مطلوب در زمینه انتخاب تاکسونومی سوالات، رعایت قواعد ساختاری، روایی محتوا و سایر شاخص های استانداردسازی آزمون برخوردار نباشد، نه تنها نقش اصلی آزمون از دست می رود بلکه این امر بر انگیزه فراگیران اثرات منفی بر

محاسبه نشد. ولی در مورد همه بخش‌های شرکت‌کننده در مطالعه (چه دسترسی به نمونه سوالات امتحانی، چه برگه‌های تصحیح شده) ارتباط و پوشش محتوایی، سطح شناختی (تاکسونومی) و میزان رعایت شدن قواعد ساختاری سوالات و قواعد نگارشی آزمون‌ها بررسی شد. بخش اندودانتیکس و دندانپزشکی جامعه نگر نیز در این تحقیق شرکت نکردند.

برای محاسبه ضریب دشواری سؤال کافی است که تعداد کل افرادی را که به آن سؤال جواب درست داده‌اند بر تعداد کل آزمون شوندگان تقسیم کنیم و نتیجه را در ۱۰۰ ضرب نماییم. هر اندازه ضریب دشواری یک سؤال بزرگ‌تر باشد، آن سؤال آسان‌تر است و هر اندازه کوچک‌تر باشد، سؤال دشوارتر است. به طور کلی ضرایب دشواری بین ۰/۳۰ تا ۰/۷۰ حداکثر اطلاع را درباره تفاوت بین آزمودنی‌ها به دست می‌دهد.

برای محاسبه ضریب تمیز یک سؤال تعداد انتخاب‌های درست گروه پایین از تعداد انتخاب‌های درست گروه بالا کسر می‌گردد و حاصل بر تعداد افراد یک گروه تقسیم می‌شود.

برای محاسبه ارزش مورد نظر برای هر گزینه انحرافی تعداد انتخاب‌های گروه بالا با تعداد انتخاب‌های گروه پایین جمع و بر مجموع تعداد دو گروه تقسیم می‌شود.

برای به دست آوردن تعداد افراد هر گروه (بالا و پایین) برای هر آزمون، لیست نمرات دانشجویان از نمرات بیشتر به کمتر مرتب شد. یک سوم تعداد دانشجویانی که نمرات بالای دامنه را کسب نموده بودند گروه بالا و یک سوم دانشجویانی که نمرات پایین دامنه را کسب کرده بودند گروه پایین را تشکیل دادند (مثلاً اگر آزمونی ۴۵ نفر شرکت‌کننده داشت ۱۵ نفر گروه بالا و ۱۵ نفر گروه پایین را تشکیل دادند). چنانچه تعداد دانشجویان بیش از ۵۰ نفر

بخش را شناسایی کرد ما برآن شدیم تا با انجام این مطالعه جامع میزان انطباق سوالات آزمون‌های نظری بخش‌های مختلف دانشکده دندانپزشکی شهیدبهشتی را با ۳۰ شاخص مربوط به اصول طراحی سؤال و آزمون چون اعتبار ساختاری، ارتباط و پوشش محتوایی (Content relevance and content coverage)، تاکسونومی، ضریب دشواری، ضریب تمیز، رعایت قواعد ساختاری، نگارش صحیح و صریح ارزیابی نماییم.

مواد و روش‌ها

در این مطالعه توصیفی- مقطعی تمامی سوالات آزمون‌های کتبی دروس ۱ نظری پایان ترم بخش‌های ارتودنسی، بیماری‌های دهان، فک و صورت، پریدانتیکس، اطفال، ترمیمی، رادیولوژی دهان، فک و صورت، جراحی دهان، فک و صورت، آسیب شناسی دهان، فک و صورت، پروتز ثابت و متحرک دانشکده دندانپزشکی شهید بهشتی در نیمسال اول سال تحصیلی ۹۱-۱۳۹۰ به صورت سرشماری از نظر ۲۴ شاخص مربوط به رعایت قواعد ساختاری طراحی سؤال و قواعد نگارشی آزمون (اعتبار صوری) و ۶ معیار مربوط به محتوای آزمون شامل پوشش محتوایی (Content coverage) یا همسویی با طرح درس، ارتباط محتوایی (Content relevance)، تاکسونومی، ضریب دشواری، ضریب تمیز، و ارزش مورد نظر برای هر گزینه انحرافی مورد ارزیابی قرار گرفتند.

با توجه به اینکه بعضی بخش‌ها (ارتودنسی، آسیب شناسی دهان، فک و صورت و پروتز ثابت و متحرک) فقط نمونه سؤال آزمون پایان ترم خود را در اختیار محقق گذاشتند و محقق دسترسی به برگه‌های تصحیح شده نداشت، در نتیجه برای این آزمون‌ها، ضریب دشواری و ضریب تمیز و ارزش مورد نظر برای هر گزینه انحرافی

باشد گروه بالا و پایین به صورت ۲۷ درصد کل دانشجویان محاسبه می‌شود.

لیست نمرات هر آزمون از طریق آموزش در اختیار محقق قرار داده شد که از لیست نمرات اصلی کپی گرفته شده اسامی دانشجویان از لیست نمرات جدا شد و فقط نمرات در لیست‌های کپی شده موجود بود. سپس برگه‌های امتحانی افراد گروه بالا و گروه پایین هر آزمون از بین برگه‌های آزمون جدا شد و برای هر سوال به صورت مجزا تعداد انتخاب‌های درست در هر گروه شمرده شد و در فرمول‌های فوق گذاشته شد تا ضریب تمیز یا ضریب دشواری برای هر سوال به دست آید.

سنجش انتخاب‌های درست برای سوالات چند گزینه‌ای یا صحیح و غلط به راحتی صورت می‌گیرد.

در سوالات تشریحی تک قسمتی نیز، مانند سوالات چند گزینه‌ای می‌توان به راحتی انتخاب‌های درست را شمرد. در سوالات تشریحی چند قسمتی برای هر قسمت به صورت جداگانه تعداد انتخاب‌های درست شمرده شد (مثلاً اگر در سوال تشریحی ذکر شده ۴ مورد را ذکر کنید برای هر مورد جداگانه تعداد انتخاب‌های درست تعیین شد).

تطابق سوالات و آزمون‌ها با قواعد ساختاری سوالات و قواعد نگارشی آزمون بر اساس چک لیست Millman سنجیده شد.^(۶)

سطح شناختی (تاکسونومی) سوالات و ارتباط و پوشش محتوایی آزمون‌ها توسط سه نفر از اعضای هیئت علمی هر بخش سنجیده شد. تاکسونومی به دو دسته یک و دو تقسیم شد. عدد ۱ برای سوالاتی که محفوظات دانشجو را می‌سنجد و فقط جنبه یادآوری را دارند و عدد ۲ برای سوالاتی که دانشجو برای پاسخگویی نیاز به تجزیه و تحلیل دارد.

ارتباط محتوایی، مشخص‌کننده مرتبط بودن سوالات آزمون با اهداف آموزشی مورد نظر آن درس می‌باشد و وقتی ۱۰۰ درصد می‌باشد که همه سوالات مرتبط با اهداف آموزشی باشند و سوالی نباشد که هدفی را نسنجیده باشد یا غیرمرتبط با اهداف یا محتوای درس مورد نظر باشد. درصد ارتباط محتوایی از تقسیم کردن تعداد سوالات مرتبط با اهداف درس بر تعداد کل سوالات و ضرب حاصل آن در ۱۰۰ به دست می‌آید.

پوشش محتوایی، مشخص‌کننده پوشش اهداف آموزشی می‌باشد یا به عبارت دیگر برای هر هدف متناسب با اهمیت آن تعداد مناسبی سوال در نظر گرفته شده باشد و هدفی نباشد که برای سنجش آن سوالی در نظر گرفته نشده باشد. درصد پوشش محتوایی از تقسیم کردن تعداد اهداف لحاظ شده در سوالات بر تعداد کل اهداف آموزشی و ضرب حاصل مربوطه در ۱۰۰ به دست می‌آید.

لازم به ذکر است که سوالات از بخش‌های مختلف دانشکده بدون نام اساتید طرح‌کننده دریافت شد و به کلیه اساتیدی که داوطلبانه و با تایید مدیران گروه برای همکاری در این تحقیق اعلام آمادگی نمودند گواهی شرکت در کارگاه مربوطه اعطا گردید.

داده‌های به دست آمده از مطالعه با استفاده از نرم‌افزار آماری SPSS با ویرایش ۱۵ مورد تجزیه و تحلیل قرار گرفت و جهت توصیف یافته‌ها از جداول فراوانی و درصد فراوانی استفاده شد.

یافته‌ها

از مجموع ۱۲ بخش تخصصی دانشکده دندانپزشکی شهید بهشتی، ۱۰ بخش با در اختیار گذاشتن سوالات پایان ترم نیمسال اول تحصیلی ۹۰-۹۱ در این مطالعه شرکت کردند. در مجموع ۵۴۴ سوال، شامل ۳۶۸ سوال

در جدول ۲ رعایت شدن ۲۴ معیار از قواعد ساختاری سوالات آزمون‌ها به تفکیک نشان داده شده است. طبق این جدول اکثر قواعد در بیش از ۹۰ درصد موارد رعایت شده‌اند و تنها «برجسته کردن کلمات منفی در سوال» در ۸۸ درصد موارد رعایت شده بود.

میزان رعایت شدن شاخص‌های کیفیت آزمون به تفکیک بخش‌های شرکت‌کننده در مطالعه، در جدول ۳ آورده شده است. این جدول نشان می‌دهد آزمون‌های دانشکده دندانپزشکی شهید بهشتی دارای ارتباط محتوایی ۸۵/۴ درصد و پوشش محتوایی ۸۴/۵ درصد بودند. همچنین قواعد ساختاری سوال در ۹۶/۴ درصد موارد و قواعد نگارشی آزمون در ۷۲/۲ درصد موارد رعایت شده بود. ضریب دشواری سوالات ۶۷/۸ درصد ضریب تمیز ۲۲/۳ درصد و ارزش مورد نظر برای هر گزینه انحرافی ۲۴/۵ درصد برآورد شد. ۸۴/۶ درصد سوالات دارای تاکسونومی ۱ و ۱۵/۴ درصد سوالات، دارای تاکسونومی ۲ بودند.

(۶۷/۵ درصد) چندگزینه‌ای، ۷۰ سوال (۱۳ درصد) تشریحی و ۱۰۶ سوال (۱۹/۵ درصد) از نوع صحیح و غلط از لحاظ ۲۴ شاخص مربوط به قواعد ساختاری سوالات و قواعد نگارشی آزمون و ۶ معیار مربوط به محتوای آزمون مورد بررسی قرار گرفتند. میزان رعایت اصول طراحی سوال در آزمون‌های بخش‌های مختلف دانشکده دندانپزشکی شهید بهشتی در جدول ۱ نشان داده شده است. بر اساس این جدول میزان رعایت قوانین ساختاری سوالات از ۹۲ درصد تا ۹۸ درصد در بخش‌های مختلف متغیر بود. همچنین میزان رعایت قواعد نگارشی آزمون‌ها از ۵۰ درصد تا ۱۰۰ درصد و درصد ضریب دشواری آزمون‌ها از ۶۰ درصد تا ۷۶ درصد متغیر بود. نتایج جدول بالا نشان می‌دهد که ضریب تمیز اکثریت آزمون‌ها ضعیف بوده است چون هیچکدام از آنها ضریب تمیز بالای ۳۰ درصد نداشتند. این درحالی است که ضریب تمیز هر چقدر به ۱ نزدیک‌تر باشد بهتر است. ولی ضریب دشواری اکثر آزمون‌ها مناسب بود چون همه آنها بین ۰/۲۵ تا ۰/۷۵ بودند که از نظر آموزشی مناسب می‌باشد.

نام آزمیون	تعداد کل	تعداد سوالات	میزان رعایت	میزان رعایت قواعد	پوشش معنایی	ارتباط معنایی	مورد ناکسونومی I	مورد ناکسونومی II	ضرب	ضرب	لرزش
آزمون	کل	به تفکیک نوع	قواعد نگارشی	ساختاری سوالات	***	***			تغییر	دشواری	مورد نظر
سوالات	اهداف		آزمون‌ها								برای
محرک	۳۶	۵۱	۳۶	۰/۰۰	۳۶/۳۶	۵۱/۵۱	۳۶/۳۶	۳۶/۳۶	۰/۰۱	۰/۰۱	-
پایانویزی	۶۸	۲۲	۶۸	۰/۰۰	۶۸/۶۸	۲۲/۲۲	۶۸/۶۸	۶۸/۶۸	۰/۰۰	۰/۰۰	-
تربیتی	۳۴	۸۱	۵۰	۰/۰۰	۳۴/۳۴	۸۱/۸۱	۳۴/۳۴	۳۴/۳۴	۰/۰۰	۰/۰۰	-
تربیتی	۱۱۰	۸۱	۷۸	۰/۰۰	۱۱۰/۱۱۰	۸۱/۸۱	۱۱۰/۱۱۰	۱۱۰/۱۱۰	۰/۰۰	۰/۰۰	۵۸
راشدینویزی	۵۰	۵۱	۵۰	۰/۰۰	۵۰/۵۰	۵۱/۵۱	۵۰/۵۰	۵۰/۵۰	۰/۰۰	۰/۰۰	-
ادبیت	۶۰	۶۱	۶۰	۰/۰۰	۶۰/۶۰	۶۱/۶۱	۶۰/۶۰	۶۰/۶۰	۰/۰۰	۰/۰۰	-
ادبیت	۷۶	۶۱	۷۶	۰/۰۰	۷۶/۷۶	۶۱/۶۱	۷۶/۷۶	۷۶/۷۶	۰/۰۰	۰/۰۰	-
ادبیت	۶۸	۵۱	۶۸	۰/۰۰	۶۸/۶۸	۵۱/۵۱	۶۸/۶۸	۶۸/۶۸	۰/۰۰	۰/۰۰	-
ادبیت	۳۶	۵۱	۳۶	۰/۰۰	۳۶/۳۶	۵۱/۵۱	۳۶/۳۶	۳۶/۳۶	۰/۰۰	۰/۰۰	-
پیداکنشی	۷۴	۳۶	۷۴	۰/۰۰	۷۴/۷۴	۳۶/۳۶	۷۴/۷۴	۷۴/۷۴	۰/۰۰	۰/۰۰	-

نشان کسر دهنده کل شاخص رایب می‌شده و صورت کسر نشان دهنده کل شاخص رایب شده است. در آزمون‌هایی که سوالات غلط دانشه اند یک شاخص اضافه شده است. مخرج کسر نشان دهنده کل شاخص رایب می‌شده و صورت کسر نشان دهنده کل شاخص رایب شده است.

از تقسیم کردن تعداد اهداف هر درس که در سیالات گنجانده شده بر تعداد کل اهداف ضرب در ۱۰۰ به دست آمده است.

از تقسیم کردن تعداد سوالات منطبق با اهداف درس بر تعداد کل سوالات ضرب در ۱۰۰ به دست آمده است.

این شاخص ها با استفاده از فرمولهای استاندارد محاسبه شده اند.

جدول ۲: میزان رعایت شدن قواعد ساختاری در سوالات بخش‌های مختلف دانشکده دندانپزشکی شهید بهشتی

نوع سوالات	شاخص‌ها	درصد موارد رعایت شده	تعداد موارد رعایت شده	تعداد کل موارد
تشریحی	از کاربرد کلمات چه کسی و چه وقت پرهیز شود	۱۰۰/۰	۷۰	۷۰
صحیح / غلط	طول متن سوالات صحیح و غلط یکسان	۱۰۰/۰	۱۰۶	۱۰۶
چند گزینه ای	همسانی طول گزینه‌ها	۹۴/۸	۳۴۶	۳۶۸
	طولانی نبودن گزینه‌ها	۹۵/۱	۳۴۷	۳۶۸
	کلمات منفی مشخص و برجسته	۸۸/۰	۳۲۱	۳۶۸
	عدم کاربرد گزینه‌های متضاد که یکی از آنها صحیح است.	۹۷/۳	۳۵۵	۳۶۸
	عدم طراحی سوالات منفی مضاعف	۹۹/۷	۳۶۶	۳۶۸
	فقط یک جواب صحیح	۱۰۰/۰	۳۶۸	۳۶۸
	گزینه انحرافی از لحاظ طولانی بودن، ساختار، کلمات به کار رفته	۹۸/۴	۳۶۰	۳۶۸
	عدم استفاده از واژه‌های مشابه در ساقه‌ها و گزینه‌ها	۹۹/۵	۳۶۴	۳۶۸
	عدم استفاده از عبارت الف و ب یا ب و ج	۹۳/۷	۳۴۵	۳۶۸
	عدم تکرار یک عبارت در همه گزینه‌ها	۹۹/۷	۳۶۶	۳۶۸
	عدم استفاده از عبارت همه موارد یا هیچکدام	۹۵/۱	۳۵۰	۳۶۸
کلیه سوالات	عدم استفاده از قیود مطلق مثل فقط، هرگز و همیشه و کلماتی مانند غالباً، بعضی وقتها	۹۹/۰	۵۳۴	۵۴۴
	رعایت نکات دستوری	۹۶/۰	۵۲۰	۵۴۴
	عدم کاربرد جای خالی در اول یا وسط متن سوال	۹۶/۷	۵۲۴	۵۴۴
	عدم استفاده از علائم اختصاری	۹۲/۰	۵۰۰	۵۴۴
	بدون اشکال انشایی	۹۹/۲	۵۳۶	۵۴۴
	بدون غلط املایی	۹۶/۰	۵۲۰	۵۴۴
	مطلب اصلی بطور کامل در متن اصلی سوال	۹۹/۰	۵۳۴	۵۴۴
	سوال واضح و روشن	۹۸/۵	۵۳۲	۵۴۴
	ارزیابی یک هدف اختصاصی	۹۷/۷	۵۲۹	۵۴۴
	طولانی نبودن نا مرتبط متن سوال	۹۹/۰	۵۳۴	۵۴۴
	استفاده از افعال مثبت	۹۳/۵	۵۰۸	۵۴۴

جدول ۳: شاخص‌های کیفیت آزمون‌های دانشکده دندانپزشکی شهید بهشتی بر اساس معدل

نام متغیر	درصد موارد رعایت شده	تعداد کل آزمون‌ها
پوشش محتوایی	۸۴/۵	۱۰
ارتباط محتوایی	۸۵/۴	۱۰
رعایت شدن قواعد ساختاری سوالات	۹۶/۴	۱۰
رعایت قواعد نگارشی آزمون	۷۲/۲	۱۰
ضریب دشواری	۶۷/۸	۷
ضریب تمیز	۲۲/۳	۶
ارزش مورد نظر برای هر گزینه انحرافی	۲۴/۵	۶
تاکسونومی I	۸۴/۶	۱۰
تاکسونومی II	۱۵/۴	۱۰

بحث

با توجه به نتایج حاصل از تجزیه و تحلیل داده‌ها، ۶۷/۵ درصد سوالات آزمون‌ها چندگزینه‌ای، ۱۳ درصد تشریحی و ۱۹/۵ درصد صحیح و غلط بوده است. این یافته‌ها با نتایج تحقیق شکورنیا^(۳) که به این نتیجه رسید بیشتر اساتید در پایان ترم از سوالات چهار گزینه‌ای استفاده می‌کنند هماهنگ است. بر این اساس می‌توان گفت که بیشتر اساتید دانشگاه از سوالات عینی جهت سنجش پیشرفت تحصیلی دانشجویان استفاده می‌کنند و از آزمون‌های ذهنی یا تشریحی کمتر استفاده می‌شود و این امر شاید به دلیل سهولت تصحیح سوالات عینی باشد.

قواعد ساختاری بین ۵۰ تا ۱۰۰ درصد در بخش‌های مختلف رعایت شده بود. مواردی که به طور کامل رعایت

شده بودند عبارت بودند از پرهیز از کاربرد کلمات چه کسی و چه وقت، یکسان بودن طول متن سوالات صحیح و غلط و داشتن فقط یک جواب صحیح. مواردی که کمتر رعایت شده بودند عبارت بودند از مشخص و برجسته بودن کلمات منفی، عدم استفاده از علائم اختصاری، همسان بودن طول گزینه‌ها، استفاده از افعال مثبت و نداشتن غلط املائی. این نتایج با یافته‌های تحقیق واعظی^(۷) و امام جمعه و زاهدی فر^(۸) و توده زعیم^(۹) همسویی دارد. بر این اساس می‌توان گفت که به طور متوسط میزان رعایت قواعد ساختاری سوالات توسط اساتید بیشتر از ۶۵ درصد می‌باشد که این امر نشان‌دهنده ضعف در رعایت قواعد ساختاری سوالات عینی می‌باشد.

در ارزیابی ۱۰۱۳ سؤال از مجموع ارزشیابی هجده درس ۹۲/۳۸ درصد سؤالات روایی محتوایی و ۸۰/۴۵ درصد سؤالات روایی صوری داشتند، مطابقت دارد. با توجه به نتایج ذکر شده می‌توان گفت که ارتباط و پوشش محتوایی آزمون اکثریت دروس مناسب بوده است یعنی سؤالات آزمون با مطالب محتوای هر فصل همخوانی دارد.

با توجه به جدول ۱، ضریب دشواری آزمون‌ها بین ۶۰ تا ۷۶ درصد بوده است. این نتایج با یافته‌های تحقیق امام جمعه و زاهدی فر^(۸) که به این نتیجه رسیدند حدود ۶۲ درصد از سؤال‌ها آسان، ۱۹/۳ درصد متوسط و ۱۸/۷ درصد آنها مشکل طراحی شده بودند همخوانی دارد. با توجه به اینکه بهترین ضریب دشواری در دامنه ۷۰-۳۰ درصد می‌باشد می‌توان گفت که ضریب دشواری آزمون‌های برگزار شده در بیشتر دروس توسط اساتید دانشکده دندانپزشکی شهید بهشتی دارای ضریب دشواری مناسب بوده است.

همچنین میزان ضریب تمیز سؤالات بین ۱۹ تا ۲۷ درصد به دست آمد. با توجه به این که ضریب تمیز بزرگ‌تر نشان‌دهنده قوه تمیز آن سوال است و هر چقدر درصد این ضریب به عدد ۱۰۰ نزدیک‌تر باشد مناسب‌تر است. می‌توان گفت که ضریب تمیز آزمون‌های برگزار شده در بیشتر دروس مختلف توسط اساتید خیلی پایین بوده و قدرت تشخیص لازم برای دانشجویان ضعیف و قوی را نداشته است.

جدول ۱ نشان می‌دهد که ارزش مورد نظر برای هر گزینه انحرافی بین ۲۰ تا ۳۱ درصد متغیر بوده است. با توجه به اینکه ضریب حدود ۲۵ درصد در سؤالات چهار گزینه‌ای برای هر گزینه انحرافی مناسب می‌باشد.^(۲)

در ۱۰۰ درصد از سؤالات از کاربرد کلمات چه کسی و چه وقت پرهیز شده است. در سؤالات صحیح / غلط در ۱۰۰ درصد موارد طول متن سؤالات یکسان بود. در سؤالات چند گزینه‌ای اکثریت قواعد ساختاری در بیش از ۹۵ درصد سؤالات رعایت شده بود و فقط برجسته کردن و مشخص کردن کلمات منفی در ۸۸ درصد سؤالات رعایت شده بود. در مجموع در ۹۵ درصد سؤالات کلیه قواعد ساختاری یعنی عدم استفاده از قیود مطلق مثل فقط، هرگز و همیشه و کلماتی مانند غالباً، بعضی وقت‌ها، رعایت نکات دستوری، عدم کاربرد جای خالی در اول یا وسط متن سوال، عدم استفاده از علائم اختصاری، بدون اشکال اثنایی، بدون غلط املائی، مطلب اصلی به طور کامل در متن اصلی سوال، سوال واضح و روشن، ارزیابی یک هدف اختصاصی، استفاده از افعال مثبت رعایت شده بود. این نتایج با یافته‌های منصوریان^(۱۰) که به این نتیجه رسیدند در اکثر موارد مشکلات مربوط به روایی صوری سؤالات چند گزینه‌ای شامل، وجود سؤالات منفی، وجود کلمات اضافه و تکراری در گزینه‌ها، ناقص بودن تنه سوال، عدم به کارگیری سطوح مختلف حیطه شناختی و به کارگیری علایم اختصاری بوده است، همخوانی دارد. بر این اساس می‌توان گفت که اکثریت استادان دانشکده دندانپزشکی دانشگاه علوم پزشکی شهید بهشتی آشنایی خوبی با قواعد ساختاری تدوین آزمون دارند.

ارتباط محتوایی آزمون‌های مورد بررسی بین ۶۹ درصد تا ۹۹ درصد درصد و به طور میانگین ۸۵/۴ درصد و پوشش محتوایی آزمون‌ها بین ۸۱ تا ۱۰۰ درصد و به طور میانگین ۸۴/۵ درصد متغیر بوده است. این نتایج با یافته‌های تحقیق کاظمی و احسان پور^(۱۱) که نشان دادند

بنابراین ارزش مورد نظر برای هر گزینه انحرافی در سوالات دروس مختلف مورد بررسی در این تحقیق مناسب می‌باشد و گزینه‌های انحرافی آزمون‌هایی که ارزش مورد نظر آنها کمتر از ۰/۲۵ می‌باشد نیاز به بازبینی دارند.

با توجه به نتایج حاصل از تجزیه و تحلیل داده‌ها، ۶۹ درصد تا ۹۹ درصد سوالات بخش‌های مختلف دانشکده دندانپزشکی از نوع تاکسونومی I بوده است و میزان تاکسونومی II، بین ۱ تا ۳۱ درصد متغیر بوده است. این نتایج با یافته‌های تحقیق شکورنیا^(۳) که نشان داد ۹۰ درصد سوالات چندگزینه‌ای در سطح تاکسونومی I طراحی شده بود، شباهت دارد. با توجه به اینکه سوالاتی که در سطح تاکسونومی I تنظیم و طراحی می‌شوند، بیشتر سطح دانش و فهمیدن پاسخگویان یا همان مفروضات ذهنی آنها را می‌سنجد و سوالاتی که در سطح تاکسونومی II تنظیم و طراحی می‌شوند بیشتر سطح کاربرد، تجزیه و تحلیل، ترکیب و ارزشیابی سطوح شناختی Bloom را می‌سنجد؛ بر این اساس می‌توان گفت که بیشتر سوالاتی که اساتید از دانشجویان امتحان می‌گرفتند مفروضات ذهنی آنها را سنجیده بود.

با توجه به نتایج حاصل از تجزیه و تحلیل داده‌ها، در کل میزان ارتباط و پوشش محتوایی دروس به ترتیب ۸۵/۴ و ۸۴/۵ درصد، رعایت شدن قواعد ساختاری (روایی صوری) سوالات ۹۶/۴ درصد، رعایت قواعد نگارشی (روایی صوری) کل آزمون ۷۲/۲ درصد، ضریب دشواری ۶۷/۸ درصد، ضریب تمیز ۲۲/۳ درصد، ارزش مورد نظر برای هر گزینه انحرافی ۲۴/۵ درصد، تاکسونومی I، ۸۴/۶ درصد و تاکسونومی II، ۱۵/۴ درصد بوده است، بر این

اساس ارتباط و پوشش محتوایی، روایی صوری سوالات و آزمون، ارزش مورد نظر برای هر گزینه انحرافی و ضریب دشواری آزمون‌های اساتید مناسب و خوب بوده‌اند، اما ضریب تمیز، میزان سوالات مربوط به تاکسونومی I و تاکسونومی II نامناسب بوده است. این یافته‌ها با نتایج تحقیقات منصوریان^(۱۰)، حق شناس^(۵) و کاظمی و احسان پور^(۱۱)، امین^(۱۲)، شکورنیا^(۳)، امام جمعه و زاهدی فر^(۸) به نوعی هماهنگی دارد.

از جمله محدودیت‌های این مطالعه می‌توان به عدم همکاری بعضی از اساتید با محقق در زمینه ارائه سوالات یا آزمون‌های تصحیح شده و عدم همکاری در زمینه پر کردن جدول مربوط به تاکسونومی و ارتباط و پوشش محتوایی، کمبود منابع اطلاعاتی درباره موضوع این تحقیق، عدم وجود حد نصاب استاندارد برای سنجش بعضی از متغیرها (مثلاً ارتباط و پوشش محتوایی یا میزان رعایت شدن قواعد ساختاری) اشاره نمود.

پیشنهاد می‌شود پژوهش‌هایی با این موضوع در دانشکده‌های دیگر نیز انجام شود و پس از بررسی وضعیت موجود در صورت لزوم کارگاه‌هایی به منظور آموزش طراحی سوال برای اساتید برگزار شود و کیفیت سوالات قبل و پس از برگزاری کارگاه بررسی گردد.

نتیجه‌گیری

آزمون‌های کتبی دروس نظری دانشکده دندانپزشکی شهید بهشتی از لحاظ ارتباط و پوشش محتوایی، رعایت قواعد ساختاری و ضریب دشواری در وضعیت مطلوبی می‌باشند ولی از نظر طراحی گزینه‌های انحرافی و تاکسونومی نیاز به بازنگری دارند.

تشکر و قدردانی

مدیران محترم گروه‌های آموزشی و اساتیدی که در مراحل مختلف انجام این تحقیق همکاری نمودند صمیمانه قدردانی می‌شود. این تحقیق با حمایت و تحت نظارت مرکز توسعه آموزش دانشگاه علوم پزشکی شهید بهشتی و دفتر توسعه آموزش دانشکده دندانپزشکی شهید بهشتی انجام گرفته است.

این مقاله بر اساس پایان نامه شماره ۳۰۸۹ دوره دکترای عمومی دانشکده دندانپزشکی دانشگاه علوم پزشکی شهید بهشتی که توسط الهام عباس زاده و با راهنمایی دکتر مریم بهاروند و مشاوره دکتر محمد حسین‌زاده و دکتر زهرا جابری انصاری در سال تحصیلی ۹۰-۹۱ انجام شد، نوشته شده است. بدین وسیله از کلیه

منابع

1. Bazargan A. Educational Evaluation. 5th ed. Tehran: Samt Co; 2001. P. 79-91.
2. Seif A. Educational Measurement, Assessment and Evaluation. 3rd ed. Tehran: Agah Co; 2007. P. 60-132.
3. Shakoornia A, Khosravi A, Shariati A, Zarei A. Survey on multiple choice questions of faculty members of Jondi Shapor Medical University of Ahwaz. The 8th National Congress of Medical Education; 2007, Kerman, Iran.
4. Sepasi H, Atari YA. The study of psychometric characteristics of Shahid Chamran University faculty members' final test scores. J Educ Psych 2006; 12(4): 1-20.
5. Haghshenas MR, Vahidshahi K, Mahmoudi M, Shahbaznejad L, Parvinnejad N, Emadi A. Evaluation of multiple choice questions in the school of medicine, Mazandaran University of Medical Sciences, the first semester of 2007. Studies in Development of Medical Education 2009; 5(2): 119-26. (Persian)
6. Millman J. Handbook of Teacher Evaluation: National Council on Measurement in Education. 1st ed. Beverly Hills: SAGE Publications; 1981. P. 146-66.
7. Vaezi A, Hosseini M, Aghily H, Mirjalili M. The investigation of applying multiple-choice question making rules in final exams of the faculty members of Yazd Medical Shahid Sadoughi University. J Med Educ Develop 2007; 2(1): 31-8. (Persian)
8. Emamjomeh MM, Zahedifar F. Analysis of multiple choice questions on health sciences at health faculty of Qazvin University of Medical Sciences (2009). The Journal of Qazvin University of Medical Sciences 2012; 15(4): 108-10. (Persian)
9. Toodehzaiem M, Daneshkazemi A, Aghili H, Haerian A, Mahdavi F. Evaluation of face-validity of multiple choice questions in special lessons of dentistry at Shahid Sadoughi University of Medical Sciences. J Med Educ Develop 2012; 7(2): 55-65. (Persian)

10. Mansourian H, Shoja M, Salimi T, Mojahed S, Karimi H, Aghili H, et al. Evaluating the course plan status of faculty staff of Shahid Saduoghi University of Medical Sciences of Yazd in the year 2002. Iran J Med Educ 2002; 2(0): 37-8. (Persian)
11. Kazemi A, Ehsanpour S. Item analysis of core theoretical courses exams for midwifery students in Isfahan University of Medical Sciences. Iran J Med Educ 2011; 10(5): 643-50. (Persian)
12. Amin M, Shayan S, Hashemi H, Poursafa P, Ebrahimi A. Analysis of multiple choice questions based on classical test theory. Iran J Med Educ 2011; 10(5): 719-25. (Persian)